



Universidad Estatal a Distancia

Vicerrectoría Académica

Escuela de Ciencias Exactas y Naturales

## **Guía de estudio**

### **Estadística de salud I (Código 3164)**

Elaborado por Ana Lorena Solís

Colaboradora: Zaira Pacheco

2009

Edición académica:

Marcela Eduarte Rodríguez



Revisión filológica:

Maritza Quesada.

# Contenido

Presentación.....	1
Descripción del curso.....	2
Acerca de la guía de estudio.....	3
Desglose de temas.....	7
<b>1. Introducción a la estadística y su aplicación en las ciencias de la salud.....</b>	<b>7</b>
1.1 Historia de la estadística.....	7
1.2 Materiales básicos.....	7
1.3 Otros materiales de apoyo.....	8
1.4 Ejercicios prácticos.....	8
1.5 Aspectos generales del tema.....	11
1.6 Resumen del tema.....	11
<b>2. Recolección de datos.....</b>	<b>12</b>
2.1 Introducción.....	12
2.2 Materiales básicos.....	14
2.3 Otros materiales de apoyo.....	14
2.4 Ejercicios prácticos.....	15
2.5 Aspectos generales del tema.....	19
2.6 Resumen del tema.....	19
<b>3. Codificación y tabulación de datos.....</b>	<b>20</b>
3.1 Introducción.....	20
3.2 Materiales básicos.....	21
3.3 Otros materiales de apoyo.....	21
3.4 Ejercicios prácticos.....	22
3.5 Aspectos generales del tema .....	25
3.6 Resumen del tema.....	26
<b>4. Presentación de los datos.....</b>	<b>27</b>
4.1 Introducción.....	27
4.2 Materiales básicos.....	27
4.3 Otros materiales de apoyo.....	28
4.4 Ejercicios prácticos.....	28
4.5 Aspectos generales del tema.....	34
4.6 Resumen del tema.....	35
<b>5. Gráficos estadísticos.....</b>	<b>35</b>
5.1 Introducción.....	35
5.2 Materiales básicos.....	36
5.3 Otros materiales de apoyo.....	36
5.4 Ejercicios prácticos.....	36
5.5 Aspectos generales del tema.....	39
5.6 Resumen del tema.....	40
<b>6. Medidas de posición y de variabilidad.....</b>	<b>41</b>
6.1 Introducción.....	41
6.2 Materiales básicos.....	41
6.3 Otros materiales de apoyo.....	42

6.4	Ejercicios prácticos.....	42
6.5	Aspectos generales del tema.....	44
6.6	Resumen del tema.....	45
<b>7.</b>	<b>Introducción a la inferencia estadística.....</b>	<b>46</b>
7.1	Introducción.....	46
7.2	Materiales básicos.....	46
7.3	Otros materiales de apoyo.....	47
7.4	Ejercicios prácticos.....	47
7.5	Aspectos generales del tema.....	49
7.6	Resumen del tema.....	50
<b>8.</b>	<b>Pruebas de significancia.....</b>	<b>50</b>
8.1	Introducción.....	50
8.2	Materiales básicos.....	51
8.3	Otros materiales de apoyo.....	51
8.4	Ejercicios prácticos.....	52
8.5	Aspectos generales del tema.....	53
8.6	Resumen del tema.....	54
<b>9.</b>	<b>Asociación o relación estadística entre dos características.....</b>	<b>55</b>
9.1	Introducción.....	55
9.2	Materiales básicos.....	56
9.3	Otros materiales de apoyo.....	56
9.4	Ejercicios prácticos.....	56
9.5	Aspectos generales del tema.....	61
9.6	Resumen del tema.....	61
<b>10.</b>	<b>Glosario.....</b>	<b>62</b>
	Bibliografía.....	63

## *Presentación*

Desde la década de los 70 hasta la fecha, Costa Rica ha experimentado un auge en el uso de la estadística en sus actividades cotidianas; es común ver cómo empresas privadas e instituciones públicas ponen a disposición de sus trabajadores datos que apoyan la toma de decisiones. Sin embargo, el mayor problema que enfrentan algunos profesionales constituye la falta de formación estadística para poder hacer uso adecuado de la información que reciben.

Dentro del campo de la salud, esta situación no es diferente y requiere una mayor atención por parte de los profesionales en registros médicos y estadística, puesto que su rol con lleva en forma inherente el manejo estadístico apropiado de datos para la toma de decisiones, la planificación y el desarrollo de su centro de salud.

En este contexto, surge la necesidad de formarse en estadística; un profesional que desconoce el manejo de datos se encuentra en desventaja para el desempeño exitoso de sus funciones. Adicionalmente, la calidad exigida por los usuarios de los servicios y su necesidad de información de, obliga ha que, cada día, el personal de registros y sistema de información en salud esté mejor preparado para el manejo de datos y su respectivo análisis.

La presente guía de estudio tiene como propósito apoyarle en el aprendizaje de la estadística aplicada al campo de la salud, en particular estimular el pensamiento analítico que usted requiere para la ejecución de sus funciones.

## *Descripción del curso*

---

### **Propósito de la asignatura**

La asignatura de Estadística tiene como propósito brindarle una herramienta que le permita tomar ciertas decisiones en el área de la salud, con base en datos estadísticos.

La asignatura le ofrece diversas técnicas estadísticas de orden descriptivo que le facilitarán la ejecución cotidiana de su trabajo; sin embargo el curso está dirigido más hacia la aplicación de las técnicas y su correcta interpretación, de forma que el futuro profesional cuente con los elementos necesarios para garantizar un uso adecuado de la estadística en el ejercicio de sus funciones.

### **Objetivos de la asignatura**

Al finalizar el curso de Estadística de salud usted estará en capacidad de:

- Aplicar conocimientos básicos de estadística para el desarrollo de investigaciones de tipo descriptivo e inferencial.
- Aplicar los elementos básicos de la estadística en la comprensión de los fenómenos de salud.
- Emplear los elementos básicos de la recolección de datos estadísticos en el campo de la salud.
- Elaborar gráficos de datos y series estadísticas.
- Interpretar gráficas y series estadísticas utilizadas en el campo de la salud.

## Acerca de la guía de estudio

---

### 1. Propósito de la guía de estudio

La presente guía de estudio constituye una herramienta de ayuda que le brindará, en cada uno de los temas, el apoyo u orientación que recibiría, en forma directa, por parte del instructor en una modalidad presencial.

En ella encontrará orientación acerca de las lecturas propias de cada tema; se le guiará respecto a lo más relevante de cada tema, mediante ejercicios resueltos y ejercicios sugeridos para su auto evaluación. Se le ofrecerá un resumen de cada tema, un glosario de términos y una ayuda esquemática que facilitará su aprendizaje.

Adicionalmente, se le motiva a usar las herramientas de Internet, las cuales reforzarán su aprendizaje.

### 2. Materiales que se utilizarán en el curso

Los materiales básicos del curso son los siguientes:

- Moya-Meoño, L. (1986). *Introducción a la Estadística de la Salud*. San José: Editorial de la Universidad de Costa Rica.
- Gómez-Barrantes, M. (1998). *Elementos de Estadística Descriptiva*. San José: Editorial EUNED.
- Del Puerto, S. (2007). "Identificación y análisis de los errores cometidos por los alumnos en estadística descriptiva". *Revista Iberoamericana de Educación* 43(3).
- Área de Estadística, Escuela de Ciencias. Facultad de Ingeniería - Universidad de San Carlos de Guatemala

<<http://sitios.ingenieria.usac.edu.gt/estadistica/estadistica2/estadisticadescriptiva.html>>. Actualizado el 21 de agosto del 2007.

- Fistera.com, web médica independiente dedicada especialmente a profesionales de Atención Primaria y orientada a la consulta y la formación.

<<http://www.fistera.com/mbe/investiga/10descriptiva/10descriptiva.asp>>. Actualizada el 6 de marzo del 2001.

- Universidad de Antioquia. Aprende en Línea.  
<<http://aprendeonline.udea.edu.co/lms/moodle/course/view.php?id=402>>.
- Hospital Universitario Ramón y Cajal. Servicio Madrileño de Salud  
<[http://www.hrc.es/bioest/M\\_docente.html#tema2](http://www.hrc.es/bioest/M_docente.html#tema2)>.

### 3. Metodología




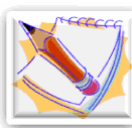

De forma sistemática usted encontrará, en cada tema, lo siguiente:






1. El capítulo o los capítulos que cubre el tema en cada uno de los libros señalados en los materiales básicos del curso.
2. Información acerca de material adicional en formatos de documentos de Word, presentaciones en Power Point o bien enlaces digitales en la web.
3. Ejercicios prácticos que refuerzan los conceptos o técnicas estadísticas tratados en el tema. Se incluyen algunas bases de datos como apoyo.
4. Indicaciones generales en relación con precaución, importancia y advertencia acerca de aspectos estadísticos que deben ser considerados para asegurar el uso correcto de la estadística.
5. Ayuda sobre el tema con los aspectos claves que el estudiante debe conocer para concluir de manera exitosa el tema, presentados en forma de resumen.
6. Una guía de ejercicios que le permitirán autoevaluar su conocimiento del tema o bien identificar si requiere reforzar algunos conceptos.



#### 4. Símbolos utilizados

Para facilitarle el manejo de esta guía, se utilizarán los siguientes símbolos que le ayudarán a identificar las actividades en forma gráfica.

Símbolo		Descripción
Información		Le indica que la actividad respectiva introduce información esencial para la comprensión del concepto o técnica estadística que se está desarrollando.
Advertencia		Significa una llamada de atención con respecto al tema que se estudia; es decir que debe poner mayor atención a las ideas señaladas, con el fin de asegurar un adecuado tratamiento de la información que está siendo expuesta al estudiante.
Precaución		Representa los aspectos que pueden llevarlo a cometer errores en el tratamiento estadístico de la información o de conceptos. Le ayuda a comprender lo que no debe hacer en el manejo de datos estadísticos.
Ejercicios		Señala que la actividad respectiva requiere del desarrollo de ejercicios de autoevaluación, los cuales le permitirán reforzar su conocimiento respecto a la técnica o concepto estadístico que se ha desarrollado en la sección correspondiente.
Autoevaluación		Se utiliza para indicarle una actividad de autoevaluación que se sugiere para reforzar los conocimientos adquiridos en el tema respectivo.

Símbolo		Descripción
Objetivos		Le indica los aspectos claves del tema tratado.
Documentos		Le refiere a documentos en Word que se recomiendan como apoyo al desarrollo del tema.
Presentaciones		Se refiere a documentos en Power Point que apoyan al desarrollo del tema.
Bases de datos		Se refiere a bases de datos relacionadas con los ejercicios que apoyan a la comprensión del tema.
Enlaces		Se refiere a enlaces recomendados como apoyo al desarrollo del tema.

## Desglose de temas

---

### 1. Introducción a la estadística y su aplicación en las Ciencias de la Salud

#### 1.1 Historia de la estadística

Cuando se habla de la estadística, generalmente se hace referencia a números, tablas, cifras demográficas, gráficos, ingresos, deudas, créditos y otros conceptos estadísticos, sin embargo, al entrar en su estudio es notorio lo fundamental que es la estadística para todas las ciencias.



Se han encontrado representaciones gráficas y símbolos en pieles, rocas y paredes de las cuevas, lo que sugiere que existían formas sencillas de estadísticas desde los inicios de la civilización. La estadística actual debe gran parte de sus logros a los trabajos matemáticos de aquellas personas que desarrollaron la teoría de las probabilidades, con la cual se adhirió la estadística a las ciencias formales.

La estadística se utiliza hoy para dar sustento científico a los datos generados de las investigaciones y los estudios de las diferentes áreas como la Psicología, las Ciencias Políticas, la Economía, la Física, la Química, la Biología, la Medicina y demás ciencias. Por lo anterior, el trabajo del experto estadístico no solo consiste en reunir y tabular los datos, sino sobre todo en interpretar esa información que se genera.

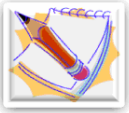
#### 1.2 Materiales básicos

Libro de texto	Capítulos que deben consultarse
Moya Meoño, L. (1986). <i>Introducción a la Estadística de la Salud</i> . San José: Editorial de la Universidad de Costa Rica.	<ul style="list-style-type: none"> <li>• Capítulo 1</li> </ul>
Gómez Barrantes, M (1998). <i>Elementos de Estadística Descriptiva</i> . San José: Editorial EUNED.	<ul style="list-style-type: none"> <li>• Tema I: Naturaleza de la Estadística: Secciones 1.1, 1.2, 1.3, 1.4, 1.9 y 1.10</li> </ul>

### 1.3 Otros materiales de apoyo

Presentaciones		Estadística de salud e investigación científica.
Enlaces		<p>&lt;<a href="http://www.fisterra.com/mbe/investiga/10descriptiva/10descriptiva.asp">http://www.fisterra.com/mbe/investiga/10descriptiva/10descriptiva.asp</a>&gt;.</p> <p>Introducción.</p> <p>Tipos de datos.</p> <p>Concepto de Estadística Descriptiva.</p>

### 1.4 Ejercicios prácticos

Ejercicios		<p>Luego de leer y estudiar su libro de texto, los materiales básicos, los materiales de apoyo y los enlaces sugeridos en esta guía, realice los siguientes ejercicios.</p>
------------	-------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------

#### Ejercicio 1: Frecuencia cardiaca

Se está interesado en estudiar si el cambio en frecuencia cardiaca observado tras la realización de una prueba de ejercicio se ve afectado por la edad de los sujetos sometidos a tal prueba. Para realizar este estudio se plantea una investigación en la que se seleccionan 40 individuos, voluntarios residentes en la provincia de Heredia; el estudio se lleva a cabo durante un día específico del mes de marzo del año 2009. Todos los individuos realizan una prueba de esfuerzo de 12 minutos y se calculan los incrementos de la frecuencia cardiaca como la frecuencia tras el ejercicio menos la frecuencia antes del ejercicio.

- Defina la unidad estadística.
- Identifique al menos dos características de la unidad y clasifíquelas según tipo en variables o atributos.

### **Solución ejercicio 1. Frecuencia Cardiaca**

a) La unidad estadística corresponde al individuo voluntario residente en la provincia de Heredia en el mes de marzo del año 2009.

b)

<b>Característica</b>	<b>Tipo</b>
Frecuencia cardiaca antes del ejercicio	Variable
Frecuencia cardiaca post ejercicio	Variable
Edad	Variable
Cambio en la frecuencia cardiaca	Variable
Género	Atributo

### **Ejercicio 2: Tabaco y función respiratoria**

En un estudio realizado por White y Froeb (1980), se investiga el efecto que sobre la función respiratoria de los no fumadores puede tener el hecho de convivir con individuos fumadores. Para tal fin se valoró la función respiratoria de individuos pertenecientes a cada uno de los siguientes grupos de fumador:

- 1: No fuma
- 2: Fumador pasivo
- 3: Fumador no inhalante
- 4: Fuma 1-10 cigarrillos

5: Fuma 11-39 cigarrillos

6: Fuma más de 40 cigarrillos

La función respiratoria fue evaluada mediante la prueba flujo de expiración forzada en L/s ( litros por segundo).

La prueba fue realizada en 2100 adultos asintomáticos seleccionados de un programa de preparación física. En este estudio, los autores demostraron que la diferencia en el flujo de expiración forzada en litros por segundo, aunque fue estadísticamente significativa, fue pequeña; esto dio pie para que el fumado fuera aprobado en lugares públicos.

- a) Defina la unidad estadística.
- b) Identifique al menos dos características de la unidad y clasifíquelas según tipo en variables o atributos.


### **Solución al ejercicio 2.**

a) La unidad estadística corresponde al adulto fumador, no fumador o no expuesto a fumadores, inscritos en un programa de preparación física, en 1980 en Nueva Inglaterra.

b)

<b>Característica</b>	<b>Tipo</b>
Flujo de expiración forzada en L/s.	Variable
Grupo de fumador	Atributo


## 1.5 Aspectos generales del tema

Información	
-------------	-----------------------------------------------------------------------------------

La unidad estadística está definida como aquella persona, elemento u objeto sobre el cual recae la observación. La validez de las conclusiones estadísticas tienen su alcance de acuerdo con el tiempo y el espacio que corresponden a la unidad estadística; así por ejemplo, la tasa de fecundidad tiene vigencia de acuerdo con el tiempo y el espacio: no es lo mismo una tasa de fecundidad en un país como Suecia que en nuestro país, pues responde a condiciones y factores culturales propios de cada país.

Advertencia	
-------------	-----------------------------------------------------------------------------------

El no tener presente el tiempo y el espacio, en relación con las conclusiones derivadas de un conjunto de unidades estadísticas, puede llevar a errores en la toma de decisiones.

Precaución	
------------	------------------------------------------------------------------------------------

El error más común en la interpretación de datos estadísticos está relacionado con la omisión de la validez de las observaciones de acuerdo con el tiempo y el espacio de la unidad estadística.

Otro de los errores comunes, aunque sea un problema de redacción, es la definición de la unidad estadística en plural, pues, como su nombre lo indica, se trata de una unidad, por tanto debe estar definida en singular.

## 1.6 Objetivos del tema

Objetivos	
-----------	-------------------------------------------------------------------------------------

Al finalizar este tema debe ser capaz de:

- Definir el concepto de estadística descriptiva e inferencial, sus alcances y limitaciones.
- Describir una unidad estadística.
- Explicar las características de orden estadístico con interés observacional e investigativo.

- d) Describir algunas aplicaciones básicas de la estadística en salud.
- e) Definir la relación entre estadística e investigación científica.

## 2. Recolección de datos

### 2.1 Introducción

La aplicación de la estadística constituye la base de la investigación científica, independientemente de si los datos existen o deben ser creados mediante algún método de recolección.

Hay diferentes tipos de datos resumidos en el siguiente cuadro:

<b>Fuente o Registro</b>	<b>Tipo</b>	<b>Ejemplos</b>
Datos existentes	Periódicos	Censo de población y vivienda. Encuesta de hogares Encuesta de causas de consulta externa
	Continuos	Estadísticas vitales: nacimientos y defunciones Estadísticas sobre egresos hospitalarios
Datos no existentes		Encuesta, observación por muestreo.

Las fuentes de datos estadísticos deben ser evaluadas con el fin de valorar dos aspectos: el primero se refiere a determinar su cobertura en términos de tiempo, espacio y periodicidad; el segundo, a la calidad de su contenido la cual describe la precisión e integridad del registro.

En el primer aspecto es frecuente encontrar errores en la omisión de datos o de unidades completas, por ejemplo en los censos de afiliación



de personas al seguro de salud de la Caja Costarricense de Seguro Social (CCSS), al igual que en censos de población, se omiten personas que normalmente se encuentran en los extremos de los grupos edad, es decir niños y adultos mayores; también, no se integran miembros de la familia que en el momento del censo no cuentan con los documentos completos, por ejemplo, estudiantes y población (hombres y mujeres) en edad productiva. Este tema es esencial en los estudios por muestreo, pues la definición incorrecta de una muestra afectará la representatividad y la validez de los datos obtenidos, como se verá en tema de muestreo.

El segundo aspecto se refiere a la calidad o interpretación de los datos debido a condicionantes sociales o de los sistemas; por ejemplo, en los servicios de salud es común la declaración incorrecta del lugar de residencia, pues este determina el acceso a un determinado establecimiento de salud. En el caso de los usuarios indocumentados, el error más frecuente es la declaración del nombre o identificación, pues esto puede generarles problemas en relación con su permanencia en el país y en el cobro de los servicios recibidos. Otro ejemplo ocurre con la edad de las personas, quienes tienden a dar el dato en forma redondeada, generalmente con décadas completas (edades terminadas en cero) o décadas y media (edades terminadas en cinco); esto afecta la precisión con que estima la edad promedio.



En general, errores como estos pueden ser subsanados con una correcta planificación del estudio y la recolección de datos; además capacitar adecuadamente a los responsables del proceso disminuye de manera significativa los errores en los datos; por otra parte, la destreza de quien recopila la información, algunas veces de quien codifica y digita puede recuperar información o mejorar de forma notable la calidad del dato.

Finalmente, otro aspecto por considerar en las fuentes de datos existentes es si estas son fuentes primarias o secundarias; las fuentes primarias, como su nombre lo indica, se refieren a productores de datos de primera mano, por ejemplo, el Instituto Nacional de Estadística y Censos, el Área de Estadística de Salud de la CCSS en relación con datos sobre hospitalización de pacientes. Las secundarias se refieren a las publicaciones en libros, revistas, páginas web o anuarios que citan la fuente principal. Es esencial considerar que los datos más fidedignos provienen de fuente primarias, sin embargo, en ocasiones la accesibilidad a estos datos obliga a utilizar fuentes secundarias, razón por la cual conocer sus limitaciones es básico.

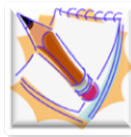
## 2.2 Materiales básicos

Libro de texto	Capítulos que deben consultarse
Moya Meoño, L. (1986). <i>Introducción a la Estadística de la Salud</i> . San José: Editorial de la Universidad de Costa Rica.	<ul style="list-style-type: none"> <li>• Capítulo 2: Planeamiento de la Recolección de Datos: Todas las secciones, excepto la 2.7.</li> <li>• Capítulo XII Problemas Prácticos que afectan la Validez de la Conclusión. Todas las secciones.</li> <li>• Capítulo XIII: Fuentes de Información en Salud. Todas las secciones.</li> </ul>
Gómez Barrantes, M. (1998) <i>Elementos de Estadística Descriptiva</i> . San José: EUNED	<ul style="list-style-type: none"> <li>• Tema II: Fuentes de Datos y Técnicas de Recolección. Todas las secciones.</li> <li>• Tema III: El Cuestionario. Todas las secciones.</li> </ul>

## 2.3 Otros materiales de apoyo

Presentaciones		Fuentes de datos existentes y no existentes en el campo de la salud
Enlaces		< <a href="http://www.fisterra.com/mbe/investiga/10descriptiva/10descriptiva.asp">http://www.fisterra.com/mbe/investiga/10descriptiva/10descriptiva.asp</a> >

## 2.4 Ejercicios prácticos

Ejercicios		Luego de leer y estudiar su libro de texto, los materiales básicos, los materiales de apoyo y los enlaces sugeridos en esta guía realice los siguientes ejercicios
------------	-------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------

## Ejercicio 1: Egresos hospitalarios de diabetes mellitus

Analice cuidadosamente el siguiente cuadro estadístico (Cuadro1), para responder a cada una de las siguientes preguntas:

**Cuadro1. Egresos hospitalarios de diabetes mellitus por Provincia de Residencia según edad y sexo del paciente. C.C.S.S. 2007**

Grupos de edad y sexo	Total	Provincia						
		San José	Alajuela	Cartago	Heredia	Guana- caste	Punta- renas	Limón
<b>Total</b>	<b>3.550</b>	<b>861</b>	<b>649</b>	<b>383</b>	<b>252</b>	<b>302</b>	<b>732</b>	<b>371</b>
<b>Total mujeres</b>	<b>1.962</b>	<b>434</b>	<b>352</b>	<b>222</b>	<b>129</b>	<b>179</b>	<b>409</b>	<b>237</b>
01 - 04	7	3	1	-	1	-	2	-
05 - 09	17	10	1	-	-	1	4	1
10 - 14	29	14	5	3	-	2	4	1
15 - 19	24	8	7	-	2	1	3	3
20 - 44	206	44	29	13	9	27	50	34
45 - 64	764	139	128	83	46	75	188	105
65 y Más	915	216	181	123	71	73	158	93
<b>Total hombres</b>	<b>1.588</b>	<b>427</b>	<b>297</b>	<b>161</b>	<b>123</b>	<b>123</b>	<b>323</b>	<b>134</b>
Menos de 1	1	-	-	-	-	-	1	-
01 - 04	7	2	2	1	1	-	1	-
05 - 09	17	3	5	1	5	-	1	2
10 - 14	23	5	10	1	-	-	2	5
15 - 19	22	10	4	4	2	-	1	1
20 - 44	231	54	44	28	12	29	50	14
45 - 64	693	171	111	78	57	68	140	68
65 y Más	594	182	121	48	46	26	127	44

FUENTE: C.C.S.S., Área de Estadística de Salud.

Pregunta	Respuesta
¿Cuál es la unidad estadística?	
¿Cuales características se incluyen? Identifique el tipo de característica	
¿Se trata de una fuente primaria o secundaria?	
¿En qué forma se recolectan los datos?	

## Ejercicio 2: Encuesta sobre prácticas sanitarias dentales

Se realizó una encuesta con el fin de estudiar las prácticas sanitarias dentales y las actitudes de una población de adultos residentes en un barrio en San José. El estudio se realizó en el mes de enero del año 2007; se entrevistaron 300 adultos, mediante el método de entrevista personal en la vivienda.

Pregunta	Respuesta
¿Cuál es la unidad estadística?	
¿Cuáles características se incluyen? Identifique el tipo de característica	
¿Se trata de una fuente primaria o secundaria?	
¿En qué forma se recolectan los datos?	


## Solución del ejercicio 1.

Pregunta	Respuesta
¿Cuál es la unidad estadística?	La unidad estadística la constituye el paciente egresado (hospitalizado) por diabetes mellitus, en los establecimientos de la CCSS en Costa Rica durante el año 2007
¿Cuales características se incluyen? Identifique el tipo de característica	Sexo del paciente. Cualitativa (atributo) Edad del paciente. Cuantitativa (ordinal) Provincia de residencia. Cualitativa (atributo)
¿Se trata de una fuente primaria o secundaria?	Se trata de una fuente primaria: el Área de Estadística de Salud recopila la información de los egresos hospitalarios de los establecimientos de la CCSS y es responsable de su publicación.
¿En qué forma se recolectan los datos?	Se trata de un registro continuo de datos, pues los centros hospitalarios remiten mensualmente los datos a la Oficina de Estadística de Salud para el procesamiento respectivo.

## Solución del ejercicio 2.


Pregunta	Respuesta
¿Cuál es la unidad estadística?	El adulto residente en el barrio seleccionado en San José en enero del año 2007
¿Cuáles características se incluyen? Identifique el tipo de característica	<p>Algunas características sugeridas son las siguientes:</p> <p>Sexo del entrevistado (atributo)</p> <p>Edad del entrevistado (cuantitativa)</p> <p>Cuántas veces al día se cepilla los dientes (cuantitativa)</p> <p>¿Cuántas veces al año visita al odontólogo para revisar su salud dental? (cuantitativa)</p> <p>¿Qué tipo de crema dental utiliza? (atributo)</p> <p>Muchas otras pueden incluirse dependiendo del objetivo del estudio.</p>
¿Se trata de una fuente primaria o secundaria?	Se trata de una investigación por muestreo; los datos no existen.
¿En qué forma se recolectan los datos?	Los datos se recolectaron mediante entrevista personal en la vivienda.

## 2.5 Aspectos generales del tema


<b>Información</b>	
--------------------	-----------------------------------------------------------------------------------

La recolección de datos constituye una de las etapas del proceso de investigación que requiere mayor cuidado, porque los resultados de la investigación dependerán del alcance y calidad de los datos (cobertura y contenido).

La recolección de datos estarán estrechamente ligada a los objetivos de la investigación y al problema que se desea resolver.

<b>Advertencia</b>	
--------------------	-----------------------------------------------------------------------------------

La elaboración de una matriz, que relacione las características por estudiar en la unidad estadística con los objetivos de la investigación, garantiza que los datos que se están obteniendo tengan un sentido en el proceso investigativo.

<b>Precaución</b>	
-------------------	-------------------------------------------------------------------------------------

La omisión de la etapa de planeamiento, en la que se establece la relación entre las características (datos) versus objetivos, es uno de los errores más frecuentes en el proceso de investigación; en muchas ocasiones el investigador concluye el estudio dejando una cantidad de datos sin analizar por falta de planificación y tiempo en el

proceso de análisis de resultados, es mejor recolectar menos datos y asegurarse su utilidad que invertir tiempo y dinero en datos que posteriormente no podrán ser analizados.

## 2.6 Objetivos del tema

<b>Objetivos</b>	
------------------	-------------------------------------------------------------------------------------

Al finalizar este tema el estudiante debe ser capaz de:

a) Describir las técnicas de la recolección de datos existentes y no existentes.

b) Definir el rol de la unidad estadística dentro del proceso de recolección de datos.

- c) Diferenciar entre variables y atributos
- d) Explicar los aspectos más relevantes por considerar en la elaboración de un cuestionario para la recolección de datos.
- e) Analizar la importancia de la definición del problema, los objetivos y la identificación de la unidad estadística dentro del proceso de recolección de datos.

### 3. Codificación y tabulación de datos

#### **3.1 Introducción**

Cada vez más, los investigadores disponen de una mayor cantidad de datos para el desarrollo de nuevo conocimiento o para apoyar la información ya existente; esto ha llevado a una mayor necesidad de conocer sobre los tratamientos estadísticos para presentar datos de forma que faciliten su interpretación y análisis.

Dentro del campo de la investigación científica, se han definido maneras de presentar la información siguiendo lineamientos estándar que dan formalidad al trabajo investigativo; de ahí lo fundamental de que el futuro especialista en el manejo de los registros y las estadísticas de salud, conozca sobre el tema y domine sus aspectos básicos.

Pese a que el auge de la aplicación de técnicas estadísticas es cada vez mayor, los problemas que se presentan siguen siendo similares, por ejemplo, se resta notoriedad a la etapa de codificación y en consecuencia en el momento de analizar la información se presentan limitaciones por codificaciones erróneas o incongruentes con el análisis de acuerdo con el tipo de característica en estudio.

La codificación de datos es una fase del proceso metodológico de investigación, la cual tiene lugar después de efectuar los trabajos de valoración de fuentes de información, fijación de objetivos, determinación de variables, diseño de hipótesis, etcétera; sin embargo, su adecuado tratamiento se verá reflejado en la etapa de presentación de datos y análisis de resultados, momento en el cual volver atrás y corregir errores resulta costoso y en la mayor parte de los casos, imposible.




También, el instrumento para la recolección de datos merece el mayor tiempo; su revisión detallada garantizará la adecuada recolección de datos; aspectos como el tipo de preguntas, tipo de características, lenguaje de redacción y prueba piloto ayudarán en el logro del objetivo de la investigación y minimizarán errores en el proceso investigativo.


### 3.2 Materiales básicos

Libro de texto	Capítulos que deben consultarse
Moya Meoño, L. (1986). <i>Introducción a la Estadística de la Salud</i> . San Jose: Editorial de la Universidad de Costa Rica.	<ul style="list-style-type: none"> <li>• Capítulo III: Elaboración de los Datos: Principios Generales. Todas las secciones.</li> <li>• Capítulo IV: Elaboración de los Datos: Tabulación. Todas las secciones.</li> </ul>
Gómez Barrantes, M. <i>Elementos de Estadística Descriptiva</i> . San José: Editorial EUNED.	<ul style="list-style-type: none"> <li>• Tema VII: Distribuciones de Frecuencias. Todas las secciones.</li> </ul>

### 3.3 Otros materiales de apoyo

Presentaciones		Distribuciones de frecuencias: algunas aplicaciones en salud.
----------------	-------------------------------------------------------------------------------------	---------------------------------------------------------------

### 3.4 Ejercicios prácticos

<b>Ejercicios</b>		Luego de leer y estudiar su libro de texto, los materiales básicos, los materiales de apoyo y los enlaces sugeridos en esta guía, realice los siguientes ejercicios
-------------------	-----------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------

#### **Ejercicio 1: Infografía de datos estadísticos de bulimia y anorexia**

Publicado por patorosario86:

<<http://patorosario86.blogspot.com/2008/10/infografia-datos-estadisticos-de.html>>

Según la Asociación de lucha contra la Bulimia y la Anorexia, las pautas culturales han determinado que la delgadez sea sinónimo de éxito social.

Muchos jóvenes luchan para conseguir el "físico ideal" motivados por modelos, artistas o por la publicidad comercial.

Durante el mes de marzo del año 2006, en el colegio "Alcántara" de la ciudad de Talca en Chile, después de las vacaciones de verano, se observó con precaución a 27 alumnos con los siguientes síntomas de anorexia:



Dieta severa	Miedo a engordar	Hiperactividad
Uso de ropa holgada	Dieta severa	Uso de laxantes
Miedo a engordar	Dieta severa	Uso de ropa holgada
Dieta severa	Uso de ropa holgada	Dieta severa
Dieta severa	Dieta severa	Uso de ropa holgada
Hiperactividad	Uso de laxantes	Miedo a engordar
Uso de laxantes	Dieta severa	Uso de ropa holgada
Uso de laxantes	Hiperactividad	Uso de laxantes
Uso de ropa holgada	Hiperactividad	Dieta severa

- Resuma la información anterior en una de distribución de frecuencias.
- Identifique el tipo de característica

Solución:

Observe que:

- Los síntomas se ordenan de mayor frecuencia a menos frecuencia.
- Se trata de una característica cualitativa o atributo.

**Distribución de los alumnos  
según síntomas de anorexia en  
el Colegio Alcántara. Talca,  
Chile Marzo 2006**

<b>Síntoma</b>	<b>Frecuencia (n<sub>i</sub>)</b>
Dieta severa	9
Uso de Ropa holgada	6
Uso de laxantes	5
Hiperactividad	4
Miedo a engordar	3
Total	27

Fuente: Colegio Alcántara. Talca, Chile.2006.

**Ejercicio 2: Estancia hospitalaria**

Se dispone de los días de hospitalización (estancia hospitalaria) de 30 pacientes del servicio de oncología del Hospital Mayo:

17	17	18	19	18	20
20	17	18	18	19	19
21	20	21	19	18	18
19	21	20	18	17	17
21	20	20	19	20	18

- Resuma la información anterior en una de distribución de frecuencias.
- Identifique el tipo de característica

Solución:

Observe que:

- Los días de estancia se ordenan de menor a mayor para dar sentido lógico a la distribución.
- Se trata de una característica cuantitativa discreta.

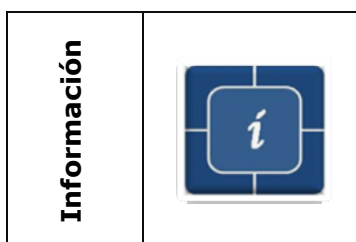
### **Distribución de los pacientes según días de estancia hospitalaria.**

#### **Hospital Mayo**


<b>Días de estancia</b>	<b>Número de pacientes (<math>n_i</math>)</b>
17	5
18	8
19	6
20	7
21	4
Total	30

Fuente: Elaboración propia.

### **3.5 Aspectos generales del tema**




El manejo de datos estadísticos requiere que el investigador, desde el principio, tenga control acerca de las características que está incluyendo en el proceso de investigación, y además conozca acerca del tipo, unidades de medida, clases o categorías de la característica; esto le ayudará a concretar con mayor facilidad la presentación de resultados y conclusiones de su investigación. Controlar estos aspectos de calidad en la medición de las características le ayudará a obtener resultados consistentes y a utilizar de manera correcta las formas de presentación de datos, según la formalidad científica requerida.

<b>Advertencia</b>	
--------------------	-----------------------------------------------------------------------------------

Asegúrese de la calidad de los datos que utilizará para respaldar una investigación: verifique que los códigos estén definidos previamente y que no se presentan ambigüedades que entorpezcan la interpretación de resultados. Utilice la presentación de datos que más se ajusta al tipo de variables y de datos; resalte aquellos aspectos que sean de interés para el lector de los


resultados, evite repetir la información; señale en todo momento las unidades de medida de la característica, lo mismo que su contexto en tiempo y espacio.

<b>Precaución</b>	
-------------------	-----------------------------------------------------------------------------------

El error más común en la presentación de datos es utilizar formas que no transmiten la idea o resultado principal. En muchos casos el investigador cae en el error de duplicar la información, y presenta cuadros y gráficos de la misma información con lo cual el lector pierde

interés en el tema. Otro error frecuente es la omisión de fuentes de los datos y sus unidades de medida; debido a esto, quien revisa la información debe recurrir a ampliarla por sus propios medios.

### 3.6 Objetivos del tema

<b>Resumen</b>	
----------------	-------------------------------------------------------------------------------------

Al finalizar este tema el estudiante debe ser capaz de:

- Utilizar los conocimientos adquiridos sobre la importancia de la etapa de codificación y tabulación de datos en el proceso de análisis estadístico.
- Aplicar una distribución de frecuencias, de acuerdo con los tipos de series estadísticas.
- Utilizar los conocimientos adquiridos para analizar los datos estadísticos mediante el uso de distribuciones de frecuencias y series estadísticas.
- Resolver ejercicios del tema, de acuerdo con los lineamientos dictados por la cátedra.

## 4. Presentación de los datos

### 4.1 Introducción

El cuadro estadístico es la forma más común de presentación de datos, aunque no la única, también lo son los gráficos; por definición un cuadro estadístico es un arreglo ordenado de los datos procesados para facilitar la lectura e interpretación de los mismos.

Para facilitar la interpretación de los datos estadísticos presentados en cuadros se han definido normas que han llevado a una forma estandarizada de presentación o criterio uniforme. El apego a las normas de construcción de cuadros facilita el análisis de la información que contiene; una presentación de datos sin estos lineamientos, afecta la claridad de la información y le resta formalidad a los trabajos de investigación.

La construcción de cuadros estadísticos es una labor aparentemente sencilla; sin embargo, es común omitir lineamientos, lo que tiene un efecto negativo sobre la validez de los datos y su confiabilidad.



Se ha comprobado que un buen estilo de presentación de cuadros y gráficos no solo facilita la interpretación del contenido, sino que lo vuelve más atractivo; al mismo tiempo reduce los costos de las publicaciones.

En el desarrollo de este tema, usted podrá encontrar recomendaciones generales que le ayudarán a mantener el criterio de uniformidad en la presentación de datos en forma de cuadros y gráficos estadísticos.


### 4.2 Materiales básicos

Libro de texto	Capítulos que deben consultarse
Moya Meoño, L. <i>Introducción a la estadística de la salud</i> . San José: Editorial de la Universidad de Costa Rica.	<ul style="list-style-type: none"> <li>Capítulo V: Presentación de datos.</li> </ul>
Gómez Barrantes, M (1998). <i>Elementos de Estadística Descriptiva</i> : Editorial EUNED.	<ul style="list-style-type: none"> <li>Tema V: Construcción y Análisis de Cuadros Estadísticos.</li> </ul>

### 4.3 Otros materiales de apoyo

Presentaciones		Cuadros Estadísticos: algunas aplicaciones en salud.
Enlaces		<a href="http://www.fisterra.com/mbe/investiga/10descriptiva/10descriptiva.asp">http://www.fisterra.com/mbe/investiga/10descriptiva/10descriptiva.asp</a> .

### 4.4 Ejercicios prácticos

Ejercicios		Luego de leer y estudiar su libro de texto, los materiales básicos, los materiales de apoyo y los enlaces sugeridos en esta guía realice los siguientes ejercicios
------------	------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------

#### Ejercicio 1: Datos sobre mamografías en la CCSS

Se dispone de los datos sobre mamografías realizadas por algunos establecimientos de la CCSS durante el periodo del año 2003 al 2006; los datos fueron suministrados por el Departamento de Estadística de los Servicios Salud.

- Elabore el título correspondiente con las preguntas básicas de la presentación de datos.
- Elabore el cuadro con las recomendaciones dadas.
- Identifique la unidad estadística, la característica de interés y el tipo.
- Interprete la información con los aspectos relevantes del cuadro.



<b>CENTRO MÉDICO</b>	<b>AÑO 2003</b>	<b>AÑO 2004</b>	<b>AÑO 2005</b>	<b>AÑO 2006</b>
H. MAX PERALTA	3.157	4.742	4.759	9.800
H. CALDERÓN GUARDIA	6.326	7.932	10.477	7.820
H. SAN JUAN DE DIOS	5.897	7.302	8.008	7.795
H. SAN VICENTE DE PAUL	4.006	4.692	6.075	6.154
CL. CORONADO	15.257	19.813	4.513	5.782
H. MÉXICO	4.773	5.938	5.037	5.430
H. ENRIQUE BALTODANO	2.924	3.883	4.216	4.846
H. DE LAS MUJERES A. CARIT	2.563	3.476	4.494	4.555
H. CIUDAD NEILLY	2.056	2.935	1.817	4.261
H. CARLOS L. VALVERDE	1.608	1.329	4.067	3.765
H. TONY FACIO	-	-	-	2.036
H. MONSEÑOR SANABRIA	1.183	1.543	1.293	1.445
H. ESCALANTE PRADILLA	797	914	668	1.046
H. SAN CARLOS	-	-	-	99
H. LOS CHILES	424	400	509	79

### **Solución:**

a) Elabore el título correspondiente con las preguntas básicas de la presentación de datos.

b) Elabore el cuadro según las recomendaciones dadas.

Observe los siguientes aspectos:

- Se eliminó el formato de celdas
- Se conservaron las líneas verticales en las columnas
- Se incluyeron los totales y la fuente.

**TOTAL DE MAMOGRAFÍAS SEGÚN EL  
CENTRO MÉDICO POR AÑO.  
C.C.S.S., 2003-2006**

<b>CENTRO MÉDICO</b>	<b>AÑO 2003</b>	<b>AÑO 2004</b>	<b>AÑO 2005</b>	<b>AÑO 2006</b>
<b><u>TOTAL</u></b>	<b><u>50.971</u></b>	<b><u>64.899</u></b>	<b><u>55.933</u></b>	<b><u>64.913</u></b>
H. MAX PERALTA	3.157	4.742	4.759	9.800
H. CALDERÓN GUARDIA	6.326	7.932	10.477	7.820
H. SAN JUAN DE DIOS	5.897	7.302	8.008	7.795
H. SAN VICENTE DE PAUL	4.006	4.692	6.075	6.154
CL. CORONADO	15.257	19.813	4.513	5.782
H. MÉXICO	4.773	5.938	5.037	5.430
H. ENRIQUE BALTODANO	2.924	3.883	4.216	4.846
H. DE LAS MUJERES A. CARIT	2.563	3.476	4.494	4.555
H. CIUDAD NEILLY	2.056	2.935	1.817	4.261
H. CARLOS L. VALVERDE	1.608	1.329	4.067	3.765
H. TONY FACIO	-	-	-	2.036
H. MONSEÑOR SANABRIA	1.183	1.543	1.293	1.445
H. ESCALANTE PRADILLA	797	914	668	1.046
H. SAN CARLOS	-	-	-	99
H. LOS CHILES	424	400	509	79

**FUENTE:** Departamento de Estadística de los Servicios  
Salud.

- c) Identifique la unidad estadística, la característica de interés y el tipo.

La unidad estadística es el procedimiento de mamografía realizado en los establecimientos de la CCSS durante el periodo 2003 al 2006.

Las características de interés y el tipo son las siguientes:

- El centro donde se aplicó el procedimientos → Cualitativa o atributo
- El año → Cuantitativa discreta

d) Interprete la información con los aspectos relevantes del cuadro.

A nivel global no se puede afirmar que exista un incremento sostenido en el número de mamografías, pues se observan variaciones importantes en los años 2004, 2005 y 2006. Llama la atención algunos hospitales que presentan cambios sobresalientes, por ejemplo, el Hospital Max Peralta duplicó el número de mamografías del 2005 al 2006; el Hospital Calderón Guardia pasó de 10 mil a 7 mil entre el 2005 y el 2006 y la Clínica de Coronado en el año 2003 presenta de 15 mil mamografías y solamente alrededor de 5 mil en el 2006.

### **Ejercicio 2: Egresos hospitalarios de *diabetes mellitus* por provincia**

Se dispone de los datos sobre egresos hospitalarios de la CCSS durante el año 2007, distribuidos por sexo, edad y provincia, los datos fueron suministrados por el Departamento de Estadística de los Servicios Salud.

- a) Elabore el título correspondiente con las preguntas básicas de la presentación de datos.
- b) Diseñe el cuadro siguiendo las recomendaciones dadas.
- c) Identifique la unidad estadística, la característica de interés y el tipo.
- d) Interprete la información con los aspectos relevantes del cuadro.

Grupos de edad y sexo	Total	Provincia						
		San José	Alajuela	Cartago	Heredia	Guana- caste	Punta- renas	Limón
Total	3.550	861	649	383	252	302	732	371
Total mujeres	1.962	434	352	222	129	179	409	237
01 - 04	7	3	1	-	1	-	2	-
05 - 09	17	10	1	-	-	1	4	1
10 - 14	29	14	5	3	-	2	4	1
15 - 19	24	8	7	-	2	1	3	3
20 - 44	206	44	29	13	9	27	50	34
45 - 64	764	139	128	83	46	75	188	105
65 y Más	915	216	181	123	71	73	158	93
Total hombres	1.588	427	297	161	123	123	323	134
Menos de 1	1	-	-	-	-	-	1	-
01 - 04	7	2	2	1	1	-	1	-
05 - 09	17	3	5	1	5	-	1	2
10 - 14	23	5	10	1	-	-	2	5
15 - 19	22	10	4	4	2	-	1	1
20 - 44	231	54	44	28	12	29	50	14
45 - 64	693	171	111	78	57	68	140	68
65 y Más	594	182	121	48	46	26	127	44

Solución:

- Elabore el título correspondiente con las preguntas básicas de la presentación de datos.
- Diseñe el cuadro siguiendo las recomendaciones dadas.

Observe los siguientes aspectos:

- Se eliminó el formato de celdas
- Se conservaron las líneas verticales en la columna matriz y de total por grupo de edad.
- Se puso en negrita los totales
- Se agregó la fuente

**Egresos Hospitalarios de Diabetes Mellitus  
por Provincia según Edad y Sexo.  
C.C.S.S. 2007**

Grupos de edad y sexo	Total	Provincia						
		San José	Alajuela	Cartago	Heredia	Guana- caste	Punta- renas	Limón
<b>Total</b>	<b>3.550</b>	<b>861</b>	<b>649</b>	<b>383</b>	<b>252</b>	<b>302</b>	<b>732</b>	<b>371</b>
<b>Total mujeres</b>	<b>1.962</b>	<b>434</b>	<b>352</b>	<b>222</b>	<b>129</b>	<b>179</b>	<b>409</b>	<b>237</b>
01 – 04	7	3	1	-	1	-	2	-
05 – 09	17	10	1	-	-	1	4	1
10 – 14	29	14	5	3	-	2	4	1
15 – 19	24	8	7	-	2	1	3	3
20 – 44	206	44	29	13	9	27	50	34
45 – 64	764	139	128	83	46	75	188	105
65 y Más	915	216	181	123	71	73	158	93
<b>Total hombres</b>	<b>1.588</b>	<b>427</b>	<b>297</b>	<b>161</b>	<b>123</b>	<b>123</b>	<b>323</b>	<b>134</b>
Menos de 1	1	-	-	-	-	-	1	-
01 – 04	7	2	2	1	1	-	1	-
05 – 09	17	3	5	1	5	-	1	2
10 – 14	23	5	10	1	-	-	2	5
15 – 19	22	10	4	4	2	-	1	1
20 – 44	231	54	44	28	12	29	50	14
45 – 64	693	171	111	78	57	68	140	68
65 y Más	594	182	121	48	46	26	127	44

FUENTE: C.C.S.S., Departamento de Estadística de Salud.

c) Identifique la unidad estadística, la característica de interés y el tipo.

La unidad estadística es el paciente egresado con el diagnóstico de *Diabetes Mellitus* en los hospitales de la CCSS en el año 2007.


Las características de interés son:


- Sexo → Cualitativa o atributo
- Provincia → Cualitativa o atributo
- Edad → Cuantitativa continua


d) Interprete la información resaltando los aspectos relevantes del cuadro.

En las provincias de San José, Alajuela y Puntarenas se presenta el mayor número de pacientes hospitalizados con el diagnóstico de *diabetes mellitus*; no se observan diferencias significativas entre hombres y mujeres pero sí entre los grupos de edad en los cuales se observa una mayor concentración en adultos; en el caso de las mujeres, la mayor concentración está en las mayores de 65 años y en el caso de los hombres, a partir de los 45 años.

#### 4.5 Aspectos generales del tema

<b>Información</b>		<p>Al elaborar cuadros estadísticos se deben tener presentes las preguntas básicas necesarias para la presentación de datos (¿Qué? ¿Cómo? ¿Cuándo? y ¿Dónde?), esto permitirá que el lector entienda fácilmente lo que el autor quiere comunicar al presentar los datos. Es recomendable tener la unidad estadística, pues facilita la elaboración del título y sus componentes. En forma adicional, seguir las recomendaciones para la elaboración de cuadros estadísticos contribuye a estandarizar en la forma de presentar los datos.</p>
--------------------	-------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

<b>Advertencia</b>		<p>La omisión de alguno de los componentes del cuadro estadístico o de respuesta a las preguntas básicas afecta la credibilidad y formalidad de la investigación, al mismo tiempo esto puede llevar a interpretaciones incorrectas de los datos y a fallas en la toma de decisiones.</p>
--------------------	-------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

<b>Precaución</b>	
-------------------	-----------------------------------------------------------------------------------

El error más común en la presentación de datos estadísticos en forma de cuadros es la elaboración del título; pocas investigaciones presentan la rigurosidad de contemplar la respuesta a las preguntas básicas y en consecuencia la interpretación de datos es parcial, incluso algunas

veces no se utiliza.

#### 4.6 Objetivos del tema

<b>Objetivos</b>	
------------------	-----------------------------------------------------------------------------------

Al finalizar este tema el estudiante debe ser capaz de:

- Identificar las formas de presentación de datos estadísticos.
- Construir cuadros estadísticos.
- Interpretar un cuadro estadístico

## 5. Gráficos estadísticos

### 5.1 Introducción

El desarrollo de herramientas computacionales ha traído una amplia gama de posibilidades en la representación gráfica; las limitaciones en el proceso de presentar datos gráficamente radican más en el usuario que en los programas mismos; lamentablemente no siempre el usuario conoce técnicamente lo que es un gráfico y, en consecuencia, se muestra la información en forma inadecuada.

La presentación gráfica de resultados es un elemento valioso para comunicar a los lectores aspectos relevantes producto del proceso de investigación; esta forma de presentación ayuda a los lectores a comprender fácilmente la idea del investigador.



Un gráfico científico debe apegarse a la realidad, no distorsionar la información proporcionada por los datos, debe comunicar ideas complejas con precisión, claridad y eficiencia, según los lineamientos dados para su elaboración; el reto se encuentra en presentarlo de manera sencilla para lograr el objetivo de comunicación.

Especial consideración debe tenerse cuando se van a comparar dos o más gráficos, pues las escalas y componentes deben ser iguales, de forma que el lector no se confunda con factores estéticos que lo hagan desviar su atención del contenido del gráfico.


## 5.2 Materiales básicos

Libro de texto	Capítulos que deben consultarse
Moya Meoño, L. <i>Introducción a la estadística de la salud</i> . San José: Editorial de la Universidad de Costa Rica.	<ul style="list-style-type: none"> <li>Capítulo 5: Presentación de datos.</li> </ul>
Gómez Barrantes, M. (1998). <i>Elementos de Estadística Descriptiva</i> : Editorial EUNED.	<ul style="list-style-type: none"> <li>Tema VI: Construcción y Análisis de Gráficos Estadísticos.</li> </ul>

## 5.3 Otros materiales de apoyo

Presentaciones		Gráficos Estadísticos: algunas aplicaciones en salud.
Enlaces		< <a href="http://www.fisterra.com/mbe/investiga/10descriptiva/10descriptiva.asp">http://www.fisterra.com/mbe/investiga/10descriptiva/10descriptiva.asp</a> >.

## 5.4 Ejercicios prácticos

Ejercicios		Luego de leer y estudiar su libro de texto, los materiales básicos, los materiales de apoyo y los enlaces sugeridos en esta guía, realice los siguientes ejercicios
------------	-------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------



### Ejercicio 1: Mujeres atendidas, según provincia.

Se dispone de la información de las mujeres atendidas en el Hospital México en el mes de agosto del año 2006, según la provincia de residencia.

a) Represente la información del cuadro en un gráfico.

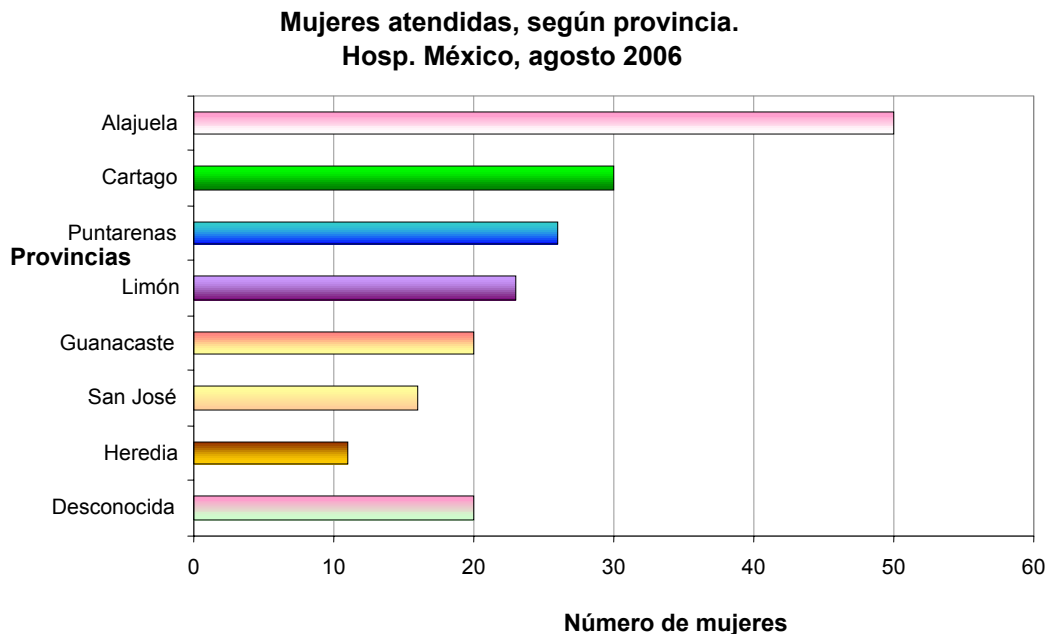
<b>Mujeres atendidas, según provincia. Hosp. México, agosto 2006</b>	
<b>Provincias</b>	<b>Total</b>
<b><u>Total</u></b>	<b><u>196</u></b>
San José	16
Alajuela	50
Cartago	30
Heredia	11
Guanacaste	20
Puntarenas	26
Limón	23
Desconocida	20

**Fuente:** Departamento de Estadística de Salud.

### Solución:

Observe lo siguiente:

- Las provincias en el gráfico fueron ordenadas de acuerdo con el número de mujeres atendidas, aunque el cuadro muestra la información en la forma convencional al orden de las provincias.



Fuente: Departamento de Estadística de Salud.

## Ejercicio 2: Mujeres atendidas, según provincia.

Se dispone de la información del peso en kilos de los pacientes egresados del servicio de Endocrinología del Hospital Dr. Raúl Blanco Cervantes, en el mes de agosto del año 1999.

a) Represente la información del cuadro en un gráfico.

Pacientes egresados del servicio de Endocrinología,  
según peso. Hosp. Dr. Raúl Blanco Cervantes,  
del mes de agosto de 1999

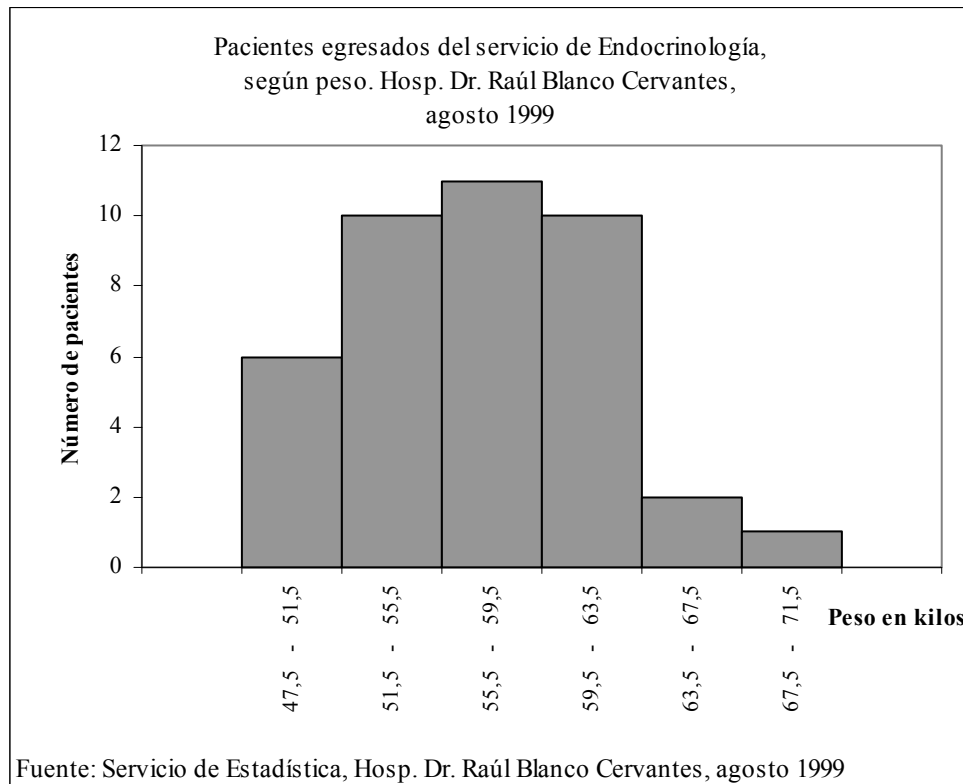
Peso en kilos	Total	Frecuencia Relativa	Porcentaje
<b>Total</b>	<b>40</b>	<b>1</b>	<b>100,00</b>
47,5 - 51,5	6	0,15	15,00
51,5 - 55,5	10	0,25	25,00
55,5 - 59,5	11	0,275	27,50
59,5 - 63,5	10	0,25	25,00
63,5 - 67,5	2	0,05	5,00
67,5 - 71,5	1	0,025	2,50

**Fuente:** Servicio de Estadística, Hosp. Dr. Raúl Blanco Cervantes, agosto 1999.

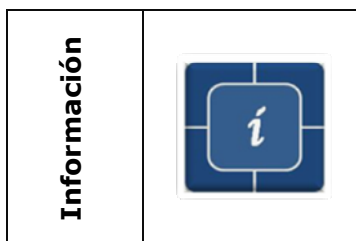
## Solución:

Observe lo siguiente:

- La distribución está representada por un histograma, cada clase está graficada al lado de la otra sin dejar espacios entre ellas.

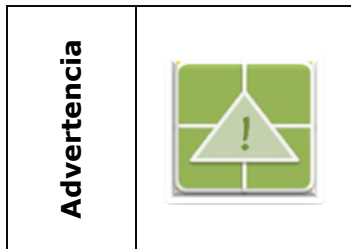


## 5.5 Aspectos generales del tema

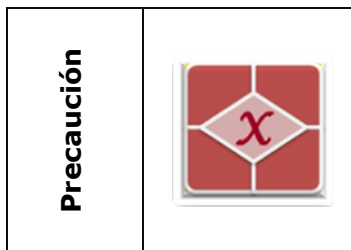


La presentación gráfica de datos debe ayudar a organizar los datos, facilitar la detección de patrones, observar agrupamientos, relaciones y la forma de la distribución de datos: esto se logra fácilmente siguiendo los lineamientos dados para la construcción de datos. Se empieza con la identificación del tipo de características y el tipo de datos, aspectos que identifican el tipo de gráfico, el cual se ajusta al requerimiento. Por otra

parte, los aspectos estéticos deben tratarse con discreción, pues el exceso puede distorsionar el objetivo básico de comunicación de un resultado mediante el gráfico.



En ocasiones, los investigadores omiten el análisis de la forma de distribución de los datos, porque les puede parecer irrelevante, sin embargo del análisis gráfico de los datos se derivan conclusiones básicas, las cuales pueden variar desde la forma hasta los conceptos de simetría y asimetría que ayudarán al análisis descriptivo, mediante alguna de las medidas de tendencia central.



El error más común en la representación gráfica de datos se refiere al abuso que existe de la gama de posibilidades que ofrecen los programas computacionales, es así como se encuentran gráficos donde se utilizan indistintamente variables continuas para presentar la información de una característica cualitativa; de manera que

no todas las opciones ofrecidas por el computador son aplicables estadísticamente.

Otro de los errores principales es la presentación de datos sin ser ordenados previamente, de manera que la idea de magnitud no llega en la primera mirada al gráfico, sino después de que el lector hace el ejercicio mental para obtener el dato más significativo.

## 5.6 Objetivos del tema



Al finalizar este tema el estudiante debe ser capaz de:

- a) Aplicar los conocimientos adquiridos para la adecuada presentación de datos estadísticos en forma gráfica.
- b) Construir gráficos estadísticos, de acuerdo con los lineamientos técnicos establecidos.
- c) Interpretar gráficos estadísticos.

## 6. Medidas de posición y de variabilidad

### 6.1 Introducción

En el campo de la salud, al igual que en los otros campos del accionar humano, la información cobra un papel relevante, la necesidad de información para la adecuada toma de decisiones impacta campos como la planificación estratégica, la compra de servicios, la gestión de las organizaciones, la dirección clínica, la satisfacción del usuario, la gestión financiera, entre otros.

Las medidas de posición y de variabilidad resumen los datos de forma que, fácilmente, se pueda conocer su distribución; hacen análisis comparativos con otros conjuntos y distinguen el grado de similitud o de concentración; esto genera, según sea el caso, información que sustente la toma de decisiones.



Los profesionales, inmersos dentro del ámbito de toma de decisiones en su establecimiento de trabajo, tienen un rol básico y protagónico: el de suministrar información en forma oportuna y con calidad.

La aplicación de las medidas de posición y de variabilidad le dará herramientas básicas de manejo estadístico de datos con la precisión, calidad y efectividad que demanda su contexto de trabajo.


### 6.2 Materiales básicos

Libro de texto	Capítulos que deben consultarse
Moya Meoño, L. <i>Introducción a la estadística de la salud</i> . San José: Editorial de la Universidad de Costa Rica.	<ul style="list-style-type: none"> <li>• Capítulo IX: Análisis e Interpretación de Datos Numéricos: Posición y Variabilidad.</li> </ul>
Gómez Barrantes, M. (1998). <i>Elementos de Estadística Descriptiva</i> : Editorial EUNED.	<ul style="list-style-type: none"> <li>• Tema VIII: Medidas de Posición.</li> <li>• Tema IX: Medidas de Variabilidad.</li> </ul>

### 6.3 Otros materiales de apoyo

<b>Presentaciones</b>		Medidas de Posición y Variabilidad: algunas aplicaciones en salud
<b>Enlaces</b>		< <a href="http://www.fisterra.com/mbe/investiga/10descriptiva/10descriptiva.asp">http://www.fisterra.com/mbe/investiga/10descriptiva/10descriptiva.asp</a> >.

### 6.4 Ejercicios prácticos

<b>Ejercicios</b>		Luego de leer y estudiar su libro de texto, los materiales básicos, los materiales de apoyo y los enlaces sugeridos en esta guía, realice los siguientes ejercicios
-------------------	-------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------

#### Ejercicio 1: Ritmo cardíaco

A un grupo de 30 personas se les ha tomado el número de pulsaciones por minuto (ritmo cardíaco), y se obtienen los siguientes resultados:

87	85	61	51	64	75	80	70	69	82
80	79	82	74	90	76	72	73	63	65
67	71	88	76	68	73	70	76	71	86

c) Calcule las medidas de posición y de variabilidad.

d) Interprete los resultados.

Solución:

Medida	Valor	Interpretación
Promedio	74,1	El número promedio de pulsaciones por minuto es de 74
Mediana	73,5	El 50% de las personas presenta un número de pulsaciones por minuto igual o menor a 74
Moda	76,0	El número de pulsaciones por minuto que se repite con más frecuencia es de 76 pulsaciones por minuto.
Varianza	78,9	No se interpreta
Desviación	8,9	En promedio, el número de pulsaciones por minuto se aparta 9 pulsaciones del promedio que es 74
Recorrido	39,0	La diferencia entre el mayor número de pulsaciones por minuto y el menor número de pulsaciones por minuto registra 39 pulsaciones.

## Ejercicio 2: Ritmo cardíaco, datos agrupados

A un grupo de 30 personas se les ha tomado el número de pulsaciones por minuto (ritmo cardíaco), y se obtiene la siguiente distribución de frecuencias:

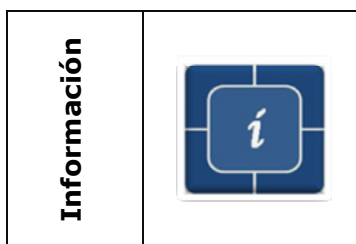
Intervalo	Punto medio	Número de personas
49,5 - 56,5	53	1
56,5 - 63,5	60	2
63,5 - 70,5	67	6
70,5 - 77,5	74	11
77,5 - 84,5	81	5
84,5 - 91,5	88	5
Total		30

- Calcule las media y la desviación estándar.
- Interprete los resultados.

### Solución:


La media es de 74,47 pulsaciones por minuto, es decir, se espera que una persona tenga en promedio 74 pulsaciones por minuto, con un desvío de 8,81 pulsaciones por minuto, **aproximadamente**.

### 6.5 Aspectos generales del tema




La adecuada aplicación de las medidas de posición o tendencia central y las medidas de variabilidad están determinados por tres factores básicamente: el tipo de características a ser estudiadas en la unidad estadística, su nivel de medición (cualitativo cuantitativa) y la forma de la distribución de datos basados en una muestra de la población. El tipo de característica y su nivel de medición indican cuál medida de posición y de variabilidad son posibles de calcular, al mismo tiempo la forma de la distribución para evaluar su condición de simetría permite evaluar cuáles medidas se ajustan técnicamente.



<b>Advertencia</b>	
--------------------	-----------------------------------------------------------------------------------

En ocasiones, se presentan resultados de análisis estadísticos inadecuados debido a un mal manejo de la codificación de las variables, así por ejemplo, si no se tiene en cuenta que se trata de una característica de tipo cualitativo y se codificó como numérica, en el momento de realizar un análisis estadístico, se calcula la media y su desviación estándar (por su condición numérica) cuando lo correcto es utilizar la moda únicamente.

<b>Precaución</b>	
-------------------	-----------------------------------------------------------------------------------

El error más común en la aplicación de medidas de posición o tendencia central se presenta en no saber escoger cuál medida describe mejor el conjunto de datos; esto ocurre cuando no se considera la condición de simetría de la distribución de los datos. Cada medida tiene sus ventajas y desventajas, la más común es obviar la condición de la media, la cual por verse afectada por valores extremos no siempre representa de la mejor manera al conjunto de datos.

Otro de los errores comunes, consiste en utilizar únicamente una medida de posición sin considerar la dispersión de la variable, con excepción de las variables ordinales; esta es una práctica incorrecta.

## 6.6 Objetivos del tema

<b>Objetivos</b>	
------------------	-------------------------------------------------------------------------------------

Al finalizar este tema el estudiante debe ser capaz de:

- 1) Definir el conceptualmente las medidas de tendencia y posición central.
- 2) Interpretar las medidas de posición y tendencia central.
- 3) Definir el concepto e interpretar las medidas de variabilidad.
- 4) Identificar los errores frecuentes en el uso de medidas de posición.
- 5) Distinguir cuáles son los errores frecuentes en el uso de medidas de posición.

- 6) Identificar algunas medidas de tendencia central utilizadas en salud y su adecuada interpretación.

## 7. Introducción a la inferencia estadística

### 7.1 Introducción

Los métodos de inferencia estadística consisten en la aplicación de un proceso de razonamiento inductivo, es decir del análisis de una parte se concluye acerca del todo; esta condición es la que se conoce como inferencia estadística.

La base de la inferencia estadística está en el muestreo, de ahí la trascendencia de que el muestreo sea representativo y aleatorio, esto por cuanto la condición de aleatoriedad permite la evaluación del error asociado a los estimadores derivados del muestreo.



La investigación científica requiere de la formalidad estadística para la generación de nuevo conocimiento o afirmación del ya existente, la rigurosidad del método permite un lenguaje estándar aprobado en todos los campos del saber humano.

En este capítulo, se aborda el concepto de inferencia estadística mediante una exposición sencilla que introduce el tema para futuras ampliaciones, por lo que se señala lo vital del muestreo y el método inductivo en la toma de decisiones.


### 7.2. Materiales básicos

Libro de texto	Capítulos que deben consultarse
Moya Meoño, L. <i>Introducción a la estadística de la salud</i> . San José: Editorial de la Universidad de Costa Rica.	<ul style="list-style-type: none"> <li>• Capítulo IX: Análisis e Interpretación de Datos Numéricos: Posición y Variabilidad.</li> </ul>
Gómez Barrantes, M. (1998). <i>Elementos de Estadística Descriptiva</i> : Editorial EUNED.	<ul style="list-style-type: none"> <li>• Tema VIII: Medidas de Posición.</li> <li>• Tema IX: Medidas de Variabilidad.</li> </ul>

### 7.3 Otros materiales de apoyo

<b>Presentaciones</b>		Introducción a la Inferencia Estadística: algunos usos en salud
<b>Enlaces</b>		< <a href="http://www.fisterra.com/mbe/investiga/10descriptiva/10descriptiva.asp">http://www.fisterra.com/mbe/investiga/10descriptiva/10descriptiva.asp</a> >.

### 7.4 Ejercicios prácticos

<b>Ejercicios</b>		Luego de leer y estudiar su libro de texto, los materiales básicos, los materiales de apoyo y los enlaces sugeridos en esta guía realice los siguientes ejercicios
-------------------	------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------

#### **Ejercicios: Selección única:**

**Marque con una "x" la opción correcta.**

1. Los índices descriptivos que definen a las muestras se denominan:

( ) Parámetros.

( ) Estimadores.

( ) Desviaciones.

( ) Distribuciones

2. Un grupo de personas afectadas por una toxoinfección alimentaria masiva, constituyen:

- ☐ Muestreo aleatorio.
- ☐ Población.
- ☐ Muestra.
- ☐ Parámetro

3. La obtención de conclusiones acerca de la población a partir de una muestra representativa es un proceso inductivo que se denomina:

- ☐ Estadística simple.
- ☐ Estadística inferencial.
- ☐ Estadística compuesta.
- ☐ Estadística descriptiva

Solución al ejercicio 1

1. Los índices descriptivos que definen a las muestras se denominan:

**(X)** Estimadores.


2. Un grupo de personas afectadas por una toxoinfección alimentaria masiva, constituyen:

**( X )** Muestra.

3. La obtención de conclusiones acerca de la población a partir de una muestra representativa es un proceso inductivo que se denomina:

**( X )** Estadística inferencial.


## 7.5 Aspectos generales del tema

<b>Información</b>	
--------------------	-----------------------------------------------------------------------------------

La necesidad de información para la toma de decisiones según el fenómeno de estudio (la vida de un paciente, el uso de un medicamento, la validez de un tratamiento, el mejoramiento de los procesos de producción de bienes y servicios, etcétera) requiere de la aplicación de técnicas estadísticas que sustente las conclusiones de


forma científica.

Siendo el muestreo la base de la inferencia estadística, es fundamental, una muestra representativa y aleatoria para la validez de las conclusiones y la toma de decisiones.

<b>Advertencia</b>	
--------------------	------------------------------------------------------------------------------------

Las muestras elaboradas sin representatividad y aleatoriedad presentan limitaciones en la medición del error de muestreo lo cual no permite la aplicación de la inferencia estadística, de manera que conservar las características deseables de una muestra es básico para la adecuada aplicación de

la inferencia.

<b>Precaución</b>	
-------------------	-------------------------------------------------------------------------------------

Los errores comunes en la aplicación de técnicas de inferencia estadística son varios, empezando con la condición de la muestra sobre la cual se basa la inferencia, hasta errores propios de las pruebas de hipótesis; por ejemplo, es común leer expresiones como: "se acepta la hipótesis nula",

"existe evidencia estadística para aceptar la hipótesis nula"; sin embargo, la existencia de los errores tipo I y tipo II nos impiden aceptar una hipótesis pero si nos permiten no rechazarla, lo correcto es concluir acerca de la posibilidad o no de rechazar la hipótesis nula conforme la evidencia estadística lo sugiera.

## 7.6 Objetivos del tema

<b>Resumen</b>	
----------------	-----------------------------------------------------------------------------------

Al finalizar este tema el estudiante debe ser capaz de:

- a) Definir el concepto de inferencia estadística
- b) Identificar las diferencias entre un parámetro y un estimado
- c) Explicar la importancia del muestreo y las características técnicas que debe tener.

## 8. Pruebas de significancia

### 8.1 Introducción

Las pruebas de significancia por excelencia constituyen una de las herramientas técnicas más potentes en el proceso de investigación científica.

Cada vez más, investigadores y científicos en general recurren al proceso de prueba de hipótesis estadística para ampliar el conocimiento sobre teorías particulares, debatir otras o incluir nuevo conocimiento.

La solidez de una prueba estadística conduce a la seguridad en la toma de decisiones, sin embargo se deben tener en consideración los riesgos que presenta esta técnica de análisis, al igual que las características deseables que reforzarán la evidencia estadística para una adecuada toma de decisiones.

En el desarrollo de este tema, se busca introducir el concepto de prueba de hipótesis y su valor técnico en la generación de conocimiento científico; la destreza en la elaboración de pruebas de hipótesis será abordada con mayor profundidad en otro curso.


## 8.2 Materiales básicos

Libro de texto	Capítulos que deben consultarse
Moya Meoño, L. <i>Introducción a la estadística de la salud</i> . San José: Editorial de la Universidad de Costa Rica.	<ul style="list-style-type: none"> <li>• Capítulo IX: Análisis e Interpretación de Datos Numéricos: Posición y Variabilidad.</li> </ul>
Gómez Barrantes, M (1998). <i>Elementos de Estadística Descriptiva</i> : Editorial EUNED.	<ul style="list-style-type: none"> <li>• Tema VIII: Medidas de Posición.</li> <li>• Tema IX: Medidas de Variabilidad.</li> </ul>

## 8.3 Otros materiales de apoyo

<b>Presentaciones</b>		Pruebas de Hipótesis Estadísticas: algunas aplicaciones en salud
-----------------------	-------------------------------------------------------------------------------------	------------------------------------------------------------------

## 8.4 Ejercicios prácticos

<b>Ejercicios</b>		Luego de leer y estudiar su libro de texto, los materiales básicos, los materiales de apoyo y los enlaces sugeridos en esta guía realice los siguientes ejercicios
-------------------	-----------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------

### Ejercicio 1:

Para cada uno de los enunciados indique falso o verdadero:

- ( ) Error tipo I: es el que se comete al rechazar la hipótesis nula ( $H_0$ ) cuando ésta es verdadera.
- ( ) Error tipo II es el que se comete al aceptar  $H_0$  cuando ésta es falsa.
- ( ) Error alfa es la probabilidad de cometer un error de tipo I.
- ( ) Error beta es la probabilidad de aceptar  $H_1$  siendo ésta verdadera

El proceso estadístico mediante el cual se desea comprobar si un valor hipotético dado a la media de la población es compatible con los datos observados en la muestra se llama

- ( ) Hipótesis múltiple.
- ( ) Hipótesis nula.
- ( ) Prueba de hipótesis.
- ( ) Distribución de t-Student.



## Solución

( V ) Error tipo I: es el que se comete al rechazar la hipótesis nula ( $H_0$ ) cuando ésta es verdadera.

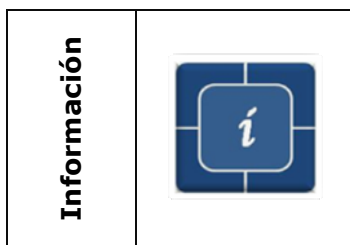
( F ) Error tipo II es el que se comete al aceptar  $H_0$  cuando ésta es falsa.

Recordar que las hipótesis estadísticas no se aceptan, la forma correcta es indicar que no existe evidencia estadística para rechazarla.

( V ) Error alfa es la probabilidad de cometer un error de tipo I.

( F ) Error beta es la probabilidad de aceptar  $H_1$  siendo ésta verdadera


### 8.5 Aspectos generales del tema




Las hipótesis estadísticas poseen características particulares, siempre se refieren a un parámetro (valor poblacional) que se supone con un valor particular o equivalente. El concepto como tal plantea dos hipótesis, las cuales son complementarias, exhaustivas y excluyentes. La hipótesis nula es la que se somete a prueba, sobre


la cual se concluye a favor o en contra; se mantiene la precaución de los errores asociados a cualquiera de las decisiones posibles. En particular, debe mantenerse especial atención en lo que la evidencia estadística plantea, lo cual debe ser concluyente en rechazar o no una hipótesis.

El proceso sobre el cual se genera la decisión acerca del supuesto valor del parámetro en estudio es lo que se conoce como prueba de hipótesis; este proceso se basa en distribuciones de probabilidad, tales como la curva normal y t-student, entre otras, para definir las zonas de rechazo y no rechazo en la cuales la evidencia estadística es concluyente.

<b>Advertencia</b>		<p>En el campo de pruebas de hipótesis estadísticas, se presentan diversos tipos de errores, entre ellos están:</p> <ol style="list-style-type: none"> <li>La calidad de la muestra estadística, la cual debe ser representativa y aleatoria.</li> <li>La hipótesis estadística debe ser clara y operacionalizable, es decir, comprensible por cualquier investigador en conceptos claramente aceptados y medibles.</li> <li>La exhaustividad de la hipótesis nula y la alternativa, debido a que en ocasiones no se da la complementariedad requerida.</li> </ol>
--------------------	-----------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

<b>Precaución</b>		<p>El error que se presenta con mayor frecuencia en el análisis de pruebas de hipótesis es la conclusión, normalmente los investigadores concluyen aceptando o rechazando la hipótesis de nulidad, lo cual es incorrecto; la hipótesis de nulidad puede rechazarse o no rechazarse, esto por que la evidencia estadística permite este alcance; la presencia de los errores tipo I y II, no permite certeza en la conclusión, por tanto, el resultado debe basarse en un nivel de confianza apoyado por la evidencia estadística para rechazar o no la hipótesis de nulidad.</p>
-------------------	-----------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

## 8.6 Objetivos del tema

<b>Objetivos</b>		<p>Al finalizar este tema el estudiante debe ser capaz de:</p> <ol style="list-style-type: none"> <li>Definir el concepto de hipótesis estadística.</li> <li>Utilizar la curva de distribución normal en el concepto de significancia estadística.</li> </ol>
------------------	-------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

## 9. Asociación o relación estadística entre dos características

### 9.1 Introducción

La técnica de asociación de variables y atributos es ampliamente utilizada en diversos campos de conocimiento, la Epidemiología, por ejemplo, se basa en técnicas de asociación estadística para valorar los factores de riesgo o bien factores protectores que potencializan el desarrollo o no de una enfermedad; en Macroeconomía, por ejemplo, a partir de cambios en los salarios puede estimarse el crecimiento de la economía, el comportamiento de los precios de los productos, etc.

Las técnicas estadísticas que estudian estas relaciones se conocen como medidas de asociación; la intensidad y dirección de esta asociación se cuantifican numéricamente mediante el uso de los coeficientes de correlación; al igual que en otras técnicas estadísticas el nivel de medición de las variables es determinante para seleccionar de forma adecuada el coeficiente que describe la asociación existente.

En el desarrollo de este tema, se estudian la medida de asociación más simple, el coeficiente de correlación de Pearson, que cuantifica la relación lineal que puede darse entre dos variables cuantitativas, por ejemplo, ingresos y gastos, peso y talla, etcétera; esto implica conocer si los valores de una de las variables tienden a ser mayores (o menores) a medida que aumentan los valores de la otra, o si no tienen relación entre sí.


La asociación entre variables es el primer paso en el estudio de causalidad, es decir en las relaciones causa efecto; no debe confundirse el alcance de las medidas de asociación con los estudios de causa efecto, para lo cual se requieren estudios de mayor rigurosidad como son los diseños experimentales y los ensayos clínicos.

Con frecuencia puede interesar la relación entre una variable y un atributo, dos atributos u otras combinaciones de niveles de medición; para cada una de estas situaciones se han definido coeficientes de correlación específicos, estos describirán con mayor precisión la relación existente; este tema será estudiado en un curso posterior.

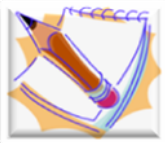
## 9.2. Materiales básicos

Libro de texto	Capítulos que deben consultarse
Moya Meoño, L. <i>Introducción a la estadística de la salud</i> . San José: Editorial de la Universidad de Costa Rica.	<ul style="list-style-type: none"> <li>Capítulo 16: Asociación o Relación Estadística entre Dos Características. Sección 16.1 y 16.2.</li> </ul>
Gómez Barrantes, M. (1998). <i>Elementos de Estadística Descriptiva</i> : Editorial EUNED.	<ul style="list-style-type: none"> <li>Tema XIV: Correlación y Regresión. Secciones 14.1-14.2-14.3 y 14.4.</li> </ul>

## 9.3 Otros materiales de apoyo

<b>Presentaciones</b>		Regresión y Correlación en dos variables: aplicaciones en el campo de la salud.
-----------------------	-------------------------------------------------------------------------------------	---------------------------------------------------------------------------------

## 9.4 Ejercicios prácticos

<b>Ejercicios</b>		Luego de leer y estudiar su libro de texto, los materiales básicos, los materiales de apoyo y los enlaces sugeridos en esta guía realice los siguientes ejercicios
-------------------	-------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------

### Ejercicio 1: Coeficiente de correlación lineal

La tabla siguiente contiene la edad X y la máxima de la presión sanguínea Y de un grupo de 10 mujeres:

Edad	56	42	72	36	63	47	55	49	38	42
Presión	14.8	12.6	15.9	11.8	14.9	13.0	15.1	14.2	11.4	14.1

- Calcule el coeficiente de correlación lineal entre las variables.
- Interprete el coeficiente en términos del problema.

Solución:

- Cálculo del coeficiente de correlación lineal:

Casos	Edad (x)	Presión (y)	$\Sigma(xy)$	$\Sigma x^2$	$\Sigma y^2$
1	56	14.8	828.8	3136	219.04
2	42	12.6	529.2	1764	158.76
3	72	15.9	1144.8	5184	252.81
4	36	11.8	424.8	1296	139.24
5	63	14.9	938.7	3969	222.01
6	47	13	611	2209	169
7	55	15.1	830.5	3025	228.01
8	49	14.2	695.8	2401	201.64
9	38	11.4	433.2	1444	129.96
10	42	14.1	592.2	1764	198.81
Total	500	137.8	7029	26192	1919.28

$$r^2 = \frac{[n \sum xy - (\sum x)(\sum y)]^2}{(n \sum x^2 - (\sum x)^2)(n \sum y^2 - (\sum y)^2)}$$

$$r^2 = \frac{[10 * 7029 - (500) * (137.8)]^2}{((10 * 26192 - (500)^2)(10 * 1919.28 - (137.8)^2))}$$

$$r^2 = \frac{1390^2}{2431203.2} = 0.79471$$

$$r = \sqrt{r^2} = \sqrt{0.79471} = 0.89146$$

- b) Interprete el coeficiente en términos del problema.

El coeficiente de correlación de Pearson tiene un valor de 0.8915, esto indica que la edad y la presión sanguínea presentan una fuerte asociación, el valor del coeficiente de correlación está cercano al valor de uno, esta asociación es positiva lo cual indica que conforme la edad aumenta es posible que la presión sanguínea también aumente.

## Ejercicio 2: Regresión Lineal

- Grafique ambas variables utilizando un grafico de dispersión.
- Determine la recta de regresión de Y sobre X.
- Interprete los coeficientes.
- Valore la bondad del modelo.
- Realice las siguientes estimaciones, señale las observaciones pertinentes:

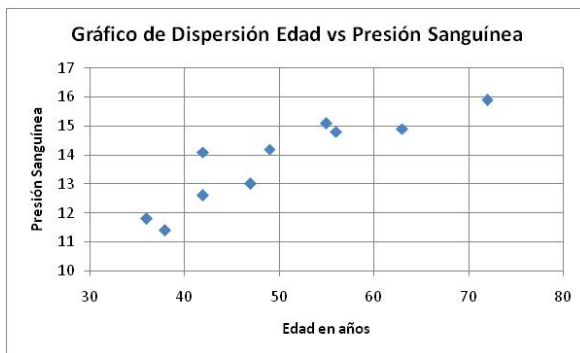
Presión sanguínea de una mujer de 51 años.

Presión sanguínea de una niña de 10 años.

Presión sanguínea de un hombre de 54 años.

## Solución

- a) Grafico de dispersión.



Como puede verse gráficamente, la relación existente entre la edad en años y la presión sanguínea es lineal creciente (positiva), por lo tanto, conforme la edad aumenta se observa un cambio en la misma dirección en la presión sanguínea.

b) Estimación de la recta de regresión de Y sobre X.

Estimación de la pendiente de la línea:

$$b = \frac{n \sum xy - (\sum x)(\sum y)}{n \sum x^2 - (\sum x)^2}$$

$$b = \frac{1390}{11920}$$

$$b = 0.1166$$

Estimación de la intercepción (a):

$$a = \bar{y} - b\bar{x}$$

$$a = 13.78 - 0.1166 * 50$$

$$a = 7.949$$

Ecuación de ajuste:

$$\begin{aligned} \hat{y} &= a + bx \\ \hat{y} &= 7.949 + 0.1166x \end{aligned}$$

La estimación de la recta de regresión indica que se trata de una recta con pendiente positiva ( $b > 0$ ).

c) Interprete los coeficientes.

La pendiente indica que se pueden esperar cambios unitarios en la edad de 0.1166 en la presión sanguínea; es decir un año más en la edad de la paciente puede implicar un aumento en la presión sanguínea en 0.1166, igualmente un año menos en la edad puede implicar una disminución de 0.1166 en la presión sanguínea.

Es importante tener en cuenta que la intercepción, pocas veces tiene sentido más allá del matemático, como punto de corte cuando la variable x asume el valor de cero, en este caso no tiene sentido, pues una edad cero está fuera del rango de estudio y no tiene sentido práctico.

d) Valore la bondad del modelo.

$$r^2 = \frac{[n \sum xy - (\sum x)(\sum y)]^2}{(n \sum x^2 - (\sum x)^2)(n \sum y^2 - (\sum y)^2)}$$

$$r^2 = \frac{[10 * 7029 - (500) * (137.8)]^2}{((10 * 26192 - (500)^2)(10 * 1919.28 - (137.8)^2))}$$

$$r^2 = \frac{1390^2}{2431203.2} = 0.79471$$

La bondad de ajuste del modelo con un valor de 79.47% indica que el modelo lineal es un buen ajuste, la relación entre la edad y la presión sanguínea, en este grupo de pacientes, se describe bastante bien con un modelo lineal. Otra interpretación puede ser orientada a la variabilidad de las variables en el modelo, así entonces podemos decir que el 79% de la variabilidad en la presión sanguínea, en el grupo de mujeres estudiadas, puede ser explicada por variaciones en la edad de la paciente.

e) Realice las siguientes estimaciones y señale las observaciones pertinentes:

*Presión sanguínea de una mujer de 51 años:*

$$\hat{y} = 7.949 + 0.1166x$$

$$\hat{y} = 7.949 + 0.1166(51) = 13.897$$

*Presión sanguínea de una niña de 10 años.*


*No procede debido a que la edad de 10 años no corresponde al rango de edad que cubre el estudio, por otra parte, se trata de una paciente demasiado diferente de la muestra en que se basa la estimación de la recta.*

*Presión sanguínea de un hombre de 54 años.*


*No procede debido a que la estimación de la recta se basa en una muestra de mujeres adultas no de hombres.*



## 9.5 Aspectos generales del tema


Información	
-------------	-----------------------------------------------------------------------------------

En el estudio de la asociación de variables y atributos es primordial tener presente el nivel de medición de cada uno de los atributos y variables, esto porque este define el tipo de coeficiente por utilizar; en el presente curso solamente se analiza la asociación entre características cualitativas.

Advertencia	
-------------	-----------------------------------------------------------------------------------

Especial atención se debe tener en el momento de aplicar estas técnicas en relación con la consistencia que deben mostrar las distintas técnicas, es decir, los resultados deben ser consistentes cuando se aplica el análisis de correlación y cuando se realizan los análisis de

asociación; el coeficiente de correlación debe mostrar consistencia con los resultados obtenidos mediante el análisis de regresión, de no ser así deben revisarse los resultados pues existe un error en alguna paso del proceso de calculo o de análisis.

Precaución	
------------	-------------------------------------------------------------------------------------

El error más frecuente en la aplicación de estas técnicas estriba en la conclusión de causalidad respecto a las relaciones existentes o descritas por los resultados, de manera que se debe tener claridad respecto al alcance de la técnica utilizada.

Otro de los errores frecuentes es la generalización de resultados fuera de los rangos de análisis, es decir, un uso inadecuado de la extrapolación de datos.

## 9.6 Objetivos del tema

Objetivos	
-----------	-------------------------------------------------------------------------------------

Al finalizar este tema el estudiante debe ser capaz de:

1) Definir el concepto de asociación estadística de dos variables cuantitativas.

2) Definir el concepto de regresión lineal simple.

3) Interpretar adecuadamente el coeficiente de correlación de Pearson.

## Glosario de conceptos

---

Término	Descripción
<b>Unidad estadística</b>	La unidad estadística está definida como aquella persona, elemento u objeto sobre el cual recae la observación. Está definida en tiempo y en espacio
<b>Variable</b>	Característica de la unidad estadística con la cual se pueden establecer diferencias cuantitativas (operaciones algebraicas) entre una clase y otra de la variable.
<b>Atributo</b>	Característica de la unidad estadística con la cual se pueden establecer diferencias cualitativas (no algebraicas) entre una categoría y otra del atributo.
<b>Egreso hospitalario</b>	Retiro de un paciente de los servicios de internamiento de un hospital.
<b>Hipótesis nula <math>H_0</math></b>	Suposición acerca de un parámetro de la población
<b>Hipótesis alternativa <math>H_1</math></b>	Una afirmación que se acepta si los datos de la muestra proporcionan evidencia acerca de que la hipótesis nula es falsa
<b>Nivel de significancia</b>	Probabilidad de rechazar la hipótesis nula cuando es verdadera
<b>Error de tipo I</b>	Rechazar la hipótesis nula cuando es verdadera
<b>Error tipo II</b>	No rechazar la hipótesis nula cuando es falsa
<b>Estadística de prueba</b>	Un valor calculado con base en la información de la muestra y se utiliza para determinar si se rechaza la hipótesis nula
<b>Valor crítico</b>	Punto de división entre la región en que se rechaza la hipótesis nula y la región en la que no se rechaza

## Bibliografía

---

Del Puerto, S (2007). Identificación y análisis de los errores cometidos por los alumnos en estadística descriptiva. *Revista Iberoamericana de Educación* 43(3).

Gómez-Barrantes, M. (1998). *Elementos de estadística descriptiva*. San José: Editorial EUNED.

Hospital Universitario Ramón y Cajal. (2009). *Material docente de la Unidad de Bioestadística Clínica*. Consultado en 31 de abril del 2009 en <[http://www.hrc.es/bioest/M\\_docente.html#tema2](http://www.hrc.es/bioest/M_docente.html#tema2)>.

Moya-Meño, L. (1986). *Introducción a la Estadística de la Salud*. San José: Editorial de la Universidad de Costa Rica.

Pita Fernández, S. (1997). Uso de la estadística y la epidemiología en atención primaria. "Manual de metodología de trabajo en atención primaria". Consultado en 31 de abril del 2009 en <<http://www.fisterra.com/mbe/investiga/10descriptiva/10descriptiva.asp>>.

S.A. 2009. *Ejercicios resueltos, Estadística descriptiva*. Consultado en 31 de abril del 2009 en <[www.inca.inf.utfsm.cl/~Erick/downloads/Cursos/Estadistica%20Computacional/EjerciciosResueltosEstadistica.pdf](http://www.inca.inf.utfsm.cl/~Erick/downloads/Cursos/Estadistica%20Computacional/EjerciciosResueltosEstadistica.pdf)>.

Universidad de Antioquia. Programa de integración de la información y la comunicación a la docencia. (2009). *Estadística descriptiva*. Consultado en 31 de marzo del 2009 en <<http://aprendeonline.udea.edu.co/lms/moodle/course/view.php?id=402>>.

Universidad de San Carlos de Guatemala. (2006). *Estadística descriptiva*. Consultado en 21 de agosto del 2007 en <<http://sitios.ingenieriausac.edu.gt/estadistica/estadistica2/estadisticadescriptiva.html>>.

